Environmetrics 00, 1–15

DOI: 10.1002/env.XXXX

Spatio-temporal functional data analysis for wireless sensor networks data

D.-J. Lee [a*], Z. Zhu [b] and P. Toscas [c]

Summary: A new methodology is proposed for the analysis, modeling and forecasting of data collected from a wireless sensor network. Our approach is considered in the framework of a functional data analysis paradigm where observed data is represented in a functional form. To reduce dimensionality, functional principal components analysis is applied to highlight important underlying characteristics and find patterns of variations. The principal scores are modeled with tensor product smooths that allow for smoothing over space and time. The model is then used for simultaneous spatial prediction at unsampled locations and to forecast future observations. We consider soil temperature data from a wireless sensor network of 50 sensor nodes in two different land types (grassland and forest) observed during 60 consecutive days in private property close to Yass, New South Wales, Australia.

Keywords: Wireless sensor networks; functional data analysis; non-parametric smoothing; penalized splines; functional principal components, forecasting

Article category: research manuscript

1. INTRODUCTION

The importance of spatio-temporal modeling has increased in recent years with the growth in large-scale data collection in various fields such as environmental monitoring applications. A wireless sensor network (WSN) is defined as a collection of battery-powered devices called

[a] *BCAM - Basque Center for Applied Mathematics, Bilbao, Spain*
[b] *Department of Statistics, Iowa State University, Ames, USA*
[c] *Commonwealth Scientific and Industrial Organization, Risk Analytics Group - Digital Productivity Flagship, Private Bag 10, South Clayton, Australia*
[*] *Correspondence to: D.-J. Lee, BCAM-Basque Center for Applied Mathematics, Mazarredo, 14, E48009 Bilbao, Basque Country, Spain. E-mail: dlee@bcamath.org, Telephone: +34 946 567 842. Fax: +34 946 567 843*

sensor nodes sending their readings in wireless radio to a base station or sink node. The data collected from WSNs allow for repeated and regular transmission and access, detect changes over the day. Throughout the days, measurements are taken at equally spaced points in time, but might not be transmitted at regular interval periods due to communication problems or energy constraints. WSNs are a very attractive technology because they are relatively cheap to deploy, relatively easy to maintain, and have minimal impact to the environment (Buratti et al., 2009; Yick et al., 2008).

This work is part of a long-term study aimed at investigating how the connection between chemical and physical properties (such as carbon, nitrogen, phosphorus and soil moisture) and microbial communities may be influenced by two different farming practices and micro-climatic conditions. As part of this work, soil grab samples will be collected a few times a year and the soils chemical, physical and microbial properties will be examined. These samples will be collected near sensor nodes that will be monitoring various climatic variables (such as ambient temperature, humidity and soil moisture) every five minutes. The ultimate aim is to examine if it is possible to relate the chemical, physical and microbial properties of the soil to the micro-climatic conditions. In this study we aim to develop methodology to interpolate soil temperature readings in space and time and to forecast readings of data from a WSN by proposing spatio-temporal functional principal components regression (Ramsay and Silverman, 2005). Since the beginning of the 1990's, functional data analysis (FDA) has been an emerging branch of Statistics for the analysis of data about curves, surfaces and other multidimensional objects, measured over a continuum. In practice, discrete data points collected over a continuum can be thought of as a function that is assumed to be a reasonably smooth mechanism giving rise to those discrete observed points. Several functional versions of a wide range of statistical tools have been developed during the last ten years (see a complete overview of FDA techniques in González-Manteiga and Vieu, 2007).

The use of WSNs data will contribute to our understanding of environmental conditions and patterns that may affect farming practices. Standard statistical techniques for FDA produce a curve over time for each location but ignores the spatial correlation between the curves and this should be accounted for. Several approaches has been proposed in the literature to deal with spatially correlated functional data extending kriging methods (Cressie, 1993). Some examples are Delicado et al. (2010), Nerini et al. (2010) or Giraldo et al. (2011). All these methods assume that the spatial functional process is stationary, i.e. no spatial trend, and isotropic, i.e. the covariance function depends on the distance between locations. However, in many practical situations, particularly in environmental sciences, the assumption of constant mean and isotropy are not realistic. Although there are several alternatives to address the challenges of (or to model) non-stationarity and anisotropy, the extension to the spatio-temporal case is a complex topic of research: first the assumption of separability between spatial and temporal covariance functions is questionable in many applications and some advances in non-separable covariance functions are at the moment too general (see Cressie and Huang, 1999 or Fuentes et al., 2007).

From a more flexible perspective, we propose a methodology for the analysis of WSNs data using low-rank smoothers such as Penalized splines (Eilers and Marx, 1996) and tensor products for space-time interpolation (Lee and Durbán, 2011), prediction and forecasting using functional data analysis, dimensionality reduction and modeling the spatial structure explicitly. The paper is organized as follows, in Section 2, we introduce the data and in Section 3, we pre-process the data using non-parametric techniques to smooth the data and interpolate the missing observations, in order to obtained a functional data representation of the measured noisy observations. Once the data are converted into smooth functional objects, we reduce the dimensionality of the data using functional principal components (FPCA) and then identify the main sources of variability. In Section 5, we use a smooth spatio-temporal model on the principal scores obtained functional principal components decomposition. Prediction and forecasting of new locations and future time points is presented in Section 6, we also evaluate the model performance by cross-validation. The paper ends with some concluding remarks.

2. THE DATA: SENSORNETS CSIRO SENSOR NETWORKS

Commonwealth Scientific and Industrial Research Organization (CSIRO) researchers are developing new technologies for outdoor sensor networks to observe physical properties of the world to a level of detail that has not been possible previously. The aim is to develop systems that can be used by scientists and other specialists to better understand the different environments. In particular the Sensornets project (http://www.sensornets.csiro.au/) developed a family of sensor network nodes known as FLECKS©, consisting of durable, and versatile devices capable of sensing, computation, actuation and wireless communications.

Our environmental monitoring example consists of a sensor network of 50 sensor nodes that are deployed in a private property close to Yass, New South Wales, Australia. These sensors collect data every 5 minutes from several environmental variables including: soil and air temperature, air humidity, soil volumetric water content and electrical conductivity. The whole study area is not very large and covers approximately 110×325 m^2 (see Figure 1). The 50 sensor nodes were divided into two land types (grassland and forest) with 25 sensor nodes on each land type. The main purpose of the study was from an engineering perspective (reliability in energy conservation and data transmission) and to understand the influence of soil in agricultural productivity and biodiversity and hence no statistical criteria on optimal designs was used to select the number of sensor nodes and their locations. We analyzed soil temperature data collected every 5 minutes intervals within a day (i.e. 288 observations/day) collected during 60 consecutive days, from 00:00 9th of November 2012 until 23:55 7th of January 2013.

[Figure 1 about here.]

3. FUNCTIONAL DATA REPRESENTATION OF SENSOR NETWORKS DATA WITH SMOOTH MODULATION P-SPLINES

In our environmental monitoring problem, each observation consists of discrete measurements $y_{t_1}, ..., y_{t_s}$ taken at time t and s location points, but these data points are assumed to arise from a (smooth) function $\mathcal{X}$, such that each $y_{t_1}, ..., y_{t_s}$ is a noisy observation of $\mathcal{X}(s,t)$. Then, the first step in a functional data analysis often consists of converting the discretely observed data to smooth functions. In a first stage, we consider each particular sensor node as an individual object, $\mathcal{X}_i(t)$, where i is the index for the i^{th} sensor node $i = 1, .., 50$, and study the characteristic of each sensor node over time ignoring the spatial distribution of the sensor nodes. We are interested in both intra-day and between-days patterns and finally among sensor nodes. For the soil temperature data on each of the 50 sensor nodes we have:

$$y_i(t) = \mathcal{X}_i(t) + \epsilon_i(t), \quad \text{for } i = 1, .., 50, \tag{1}$$

where t is time (from 00:00 of 9/11/2012 to 23:55 of 7/01/2013), and t is a vector of length 17280 (288 observations/day $\times$ 60 consecutive days). The residuals ϵ are independent with the original function $\mathcal{X}_i$ that generates the observations y_i, and $\mathbb{E}[\epsilon_i(t)] = 0$ and $\mathbb{V}\text{ar}[\epsilon_i(t)] = \sigma^2$.

While many smoothing methods are available in order to estimate a functional data from noisy observations, features of soil temperature data in our study influence which approach is most appropriate. We need a flexible structure as well as a computationally efficient method to capture the daily cyclical pattern and possibly non-stationary and non-linear trends. For times series data with cyclical patterns Eilers et al. (2008) proposed a *smooth modulation model*. Indeed, with 17820 observations per sensor node, we would require a very large number of knots to achieve a smooth fit that describes the functional data $\mathcal{X}$. Smooth modulation model allows for decomposing the data as a time trend and daily cyclical pattern using a much lower number of basis functions for smoothing in the essence of Penalized splines (P-splines) by Eilers and Marx (1996) with a combination of varying-coefficients terms (Hastie and Tibshirani, 1990) for cyclical sine a co-sine amplitudes (or modulation). The data were smoothed using a total of 165 coefficients (33 for the temporal trend, and $33 \times 2 = 66$ coefficients for the sine and co-sine modulation). The selection of the number of basis functions were based on computational efficiency and robustness. The smoothing parameters were estimated by REML (Wood, 2006). Figure 2 shows the smoothed data for a sensor node arranged by time-of-day and by days, and smoothed daily curves. Using the smooth modulation model we were able to remove the noise, interpolate the missing observations and provide a functional smooth representation of soil temperature, $\mathcal{X}_i$, for each i^{th} sensor node. The model is fitted separately to all the 50 sensor nodes.

[Figure 2 about here.]

4. FUNCTIONAL PRINCIPAL COMPONENTS ANALYSIS FOR WSNS SOIL TEMPERATURE DATA

Functional Principal Components Analysis (FPCA) is the functional extension of the multivariate principal component analysis. FPCA is a dimension reduction method, the idea behind it is to transform the sampled curves in such a way that only a low dimensional space represents the patterns of variability of the curves. General approaches are described by Ramsay and Silverman (2005) and Ferraty and Vieu (2006). We separated the estimated functional data, $\mathcal{X}(\cdot)$, by land types and reduced the dimensionality through FPCA. A more formal test to evidence the difference between land types can be done by using a functional t-test for the difference betweem grassland and forest (e.g. `fperm.fd` function in `R` library `fda`, see (Ramsay and Silverman, 2005) for details). The FPCA can be expressed as:

$$\mathcal{X}_{g,j}(t') = \mu_g(t') + \sum_{k=1}^{K} \beta_{g,j,k}\phi_{i,k}(t') + e_{g,j,k}(t'), \text{ for } j = 1, ..., 60 \text{ days, and g=1,2.} \tag{2}$$

Note that t' represent the vector of 288 daily observations and $\mu_g(t')$ is the mean for each land type ($g = 1$ for grassland and $g = 2$ for forest), $\{\phi_{g,k}(t')\}$ is a set of orthonormal basis functions (functional principal components), $\beta_{g,j,k}$ is the k^{th} principal component scores, and $e_{g,j,k}$ are *i.i.d.* residuals. The decomposition in Equation (2) is achieved using optimal orthonormal functional principal components analysis based on Ramsay and Dalzell (1991) and implemented in the `ftsa` R library (Hyndman and Shang, 2013) by the function `ftsm`. The aim is to find a set of K orthonormal functions $\{\phi_{g,k}(t')\}$, so that the expansion of each curve in terms of these basis functions approximates the curve as closely as possible. We fixed the number of basis functions to $K = 3$ because they represent almost the total proportion of the explained variability (see Table 1). In a more general setting, the number of K functions can be chosen by cross-validation (Ramsay and Silverman, 2005). For more robust estimates of the algorithm one can replace the mean $\mu_i(t')$ by the median of the smoothed curves, and also obtain robust principal components (see Hyndman and Ullah, 2007, for details).

[Table 1 about here.]

In the FPCA model in Equation (2) the 1$^{\text{st}}$ PCA component explains most of the variability (see Table 1), and it is interpreted as the deviations over the daily mean through the 60 consecutive days in both sensor nodes locations. Second and third principal components are more complex to interpret explaining much less percentage of variability as shown in Table 1, and they seem to be related with daily cyclical changes or other factors affecting soil temperature in both locations. The 2$^{\text{nd}}$ PC's are very similar for both land types, whereas the third principal component has the same pattern with opposite sign effect. Figure 3 shows functional means $\mu_1(t')$, $\mu_2(t')$, the first $K = 3$ functional principal components, $\{\phi_{i,k}(t')\}$, for the sensor nodes in grassland (in red) and forest locations (in blue).

[Figure 3 about here.]

5. SPATIO-TEMPORAL FUNCTIONAL PRINCIPAL COMPONENTS REGRESSION OF SOIL TEMPERATURE DATA

In previous section, we analyzed the main sources of variability between grassland and forest locations. From a practical point of view, the decomposition of the functional data smoothed by the smooth modulation model in Section 3 into a functional basis representation in Equation (2) reduces the problem to model the estimated basis coefficients $\hat{\beta}_{g,j,k}$. The FPCA score coefficients $\hat{\beta}_{g,j,k}$ are uncorrelated so they can be modeled independently. The next step is to incorporate the spatial location to model the principal components scores. The idea is to use a spatio-temporal model for the $\hat{\beta}_{g,j,k}$ coefficients and use them to predict a new location and forecast future values of soil temperature.

5.1. Smooth-ANOVA model for spatio-temporal data

We propose a non-separable model, where the principal scores β's from the FPCA are modeled as a sum of terms using an ANOVA decomposition of smooth functions. This model is called Smooth-ANOVA model, as analogy to classic ANOVA decomposition (Gu, 2002), where the process is represented in terms of the sum of main effects (of space and time) and interaction (space-time). Now, we incorporate the spatial information of the sensor nodes locations, hence for each sensor node we have a new index s to denote the geographic locations. The model for each k^{th} principal component scores can be formulated as:

$$\hat{\mathbf{z}}_k = f_s(location) + f_t(days) + f_{s,t}(location, days) + e, \tag{3}$$

where $\mathbf{z}_k$ is the vectorized form of the principal scores, i.e. $\text{vec}(\widehat{\beta}_{s(i),j,k})$, f_s, f_t and $f_{s,t}$ are smooth functions of space, time, and space-time interaction, and e are *i.i.d.* Gaussian errors with zero mean and variance σ^2. Grassland and forest locations are considered separately, and hence the model in Equation (3) is fitted to the first $i = 1, ..., 25$ sensor nodes in grassland and to the next 25 ($i = 26, ..., 50$) sensor nodes in the forest independently. Identifiability constraints must be considered for each the functions in (3). Lee and Durbán (2011) proposed an efficient method using tensor products of B-splines and penalties, where identifiability constraints are imposed on the regression coefficients directly. Wood (2006), implemented similar identifiability constraints in the `R` package `mgcv`, and hence model (3) can be implemented by the functions `gam` or `gamm` for the mixed model representation, with tensor product smooths denoted by `te`. Notice that the model in Equation (3) assumes a non-separable covariance structure, between space and time, as space-time interaction is explicitly considered in $f_{s,t}$ and it is a more flexible alternative for modeling multidimensional data in a non-parametric sense. Figures 4 and 5 show the fitted model in Equation (3) for the 1^{st} PC for grassland and forest locations respectively. Note that, top panels a) and b) show the main effects for space and time for fitted 1^{st} PC coefficients and bottom panels c) and d) shows the space-time interaction plotted by days, and the fitted coefficients for the 1^{st} PC respectively. Similar figures can be plotted for second and third principal components but are not shown.

[Figure 4 about here.]

[Figure 5 about here.]

6. PREDICTION AND FORECASTING OF A FUNCTIONAL DATA

Now, we are interested in predicting the value of the functional data in a new spatial location s^* and forecast new future functional data observations h-days ahead. The main idea is to use the obtained principal components scores in Equation (3) to predict new values of $\hat{\beta}_{s^*(i),j+h,k}$ in a new location s^* and h periods ahead. For illustrative purposes, let us consider first the spatial prediction of a new location s^* within grassland and forest locations. Then, we illustrate the future observations forecasting.

6.1. Spatial prediction of a new functional data

Given fitted model (3) the prediction of a new functional data in a new location is done by:

$$\widehat{\mathcal{X}}_{s^*(g),j}(t') = \widehat{\mu}_g(t') + \sum_{k=1}^{K} \widehat{\beta}_{s^*(g),j,k}\ \phi_{g,k}(t'), \tag{4}$$

where $s^*(g)$ denotes the new location in land type $g = 1, 2$, and j is the day. Observe that, $\widehat{\mu}_g(t')$ is functional mean for each land type, and the k^{th} principal component $\phi_{g,k}(t')$ obtained in Equation (2) . In this model we are considering a fixed set of principal components $\phi_{1,k}(t')$ and $\phi_{2,k}(t')$ for all the sensor nodes within the same land type. Hence, we know a priori if the new location $s^*(i)$ is located in grassland or in forest. Model (4) requires an estimate for the principal score at s^*, i.e. $\widehat{\beta}_{s^*(g),j,k}$. The principal score $\widehat{\beta}_{s^*(g),j,k}$ is estimated with the model coefficients obtained in the fitting of the smooth-ANOVA model for the PC scores. The approximated functional data at s^* is obtained by replacing the estimates in Equation (4).

6.2. Forecasting observations of a functional data

Now, we forecast the value of the principal scores coefficients $\widehat{\beta}_{s(g),j+h,k}$, h-days ahead conditioned to the fixed set of functional principal components $\phi_{g,k}(t')$, for $g = 1, 2$. Then, the estimated coefficients are used to approximate the functional data using Equation (2), i.e. the new functional data is

$$\widehat{\mathcal{X}}_{s(g),j+h}(t') = \widehat{\mu}_g(t') + \sum_{k=1}^{K} \widehat{\beta}_{s(g),j+h,k}\ \phi_{g,k}(t'), \tag{5}$$

Observe that with second order penalties linear extrapolation is used for forecasting observations h-periods ahead. Figure 6 show the forecast of the functional data $\widehat{\mathcal{X}}_{s(g),j+h}(t')$ corresponding to the sensor node ID 1 for $h = 1$ and $h = 2$ days ahead by time-of-day (t'). The forecast variance of the approximated functional data also follows from Equation (5), where because of the way the model has been constructed, each k^{th} component is orthogonal to the other components and so the forecast variance can be approximated by the simple sum of component variances.

[Figure 6 about here.]

6.3. Model performance

The performance of the proposed method is evaluated by cross-validation. We removed one sensor node on each land type, then we fit the model to the functional data up to day $60 - m$ for validation and predict the next m periods to obtain measures of the forecast accuracy. The removed sensor nodes are sensor ID 1 (in grassland) and sensor ID 28 (in forest) in Figure 4a and Figure 5a respectively. We compare the proposed method with an ordinary kriging model for the principal scores vector $\hat{\mathbf{z}}_k$ for $k = 1, 2, 3$, where coefficients are three-dimensional arrays and the model is fitted using the function `krige` in the `R` library `gstat`. An isotropic variogram model in the three dimensions was chosen in the absence of prior information on the covariance function. We choose an exponential covariance function with default values for the partial sill of 1 and range of 2 (Pedesma, 2004). To evaluate the forecasts accuracy, we computed the *out-of-sample* mean square error (MSE) and the mean absolute error (MAE). For each forecast horizon m, the measures are defined as

$$MSE_m = \frac{\sum_{t=1}^{N} E_t^2}{N} \quad \text{and} \quad MAE_m = \frac{\sum_{t=1}^{N} |E_t|}{N}, \quad \text{where } t = 1, ..., N \tag{6}$$

where $N = 288 \times (60 - m)$, E_t is the forecast error for the corresponding period $E_t = \widehat{\mathcal{X}}_{s^*(g),j,}(t') - \widetilde{\mathcal{X}}_{s^*(g),j}(t')$, with $\widehat{\mathcal{X}}_{s^*(g)}(t')$ the actual functional data value approximated in Equation (2) and $\widetilde{\mathcal{X}}_{s^*(g),j}(t')$ is the forecasted value.

The comparisons of the MSE_m and MAE_m onto the log scale are shown in Figure 7. Our approach present better results for the 1 to 5 days ahead forecasts. Forecasting performance deteriorates as we increase the forecast horizon in both methods. It is important to remark that we used a generic spatio-temporal ordinary kriging using a standard available software in the comparison with a separable isotropic exponential covariance function for space and time.

[Figure 7 about here.]

7. CONCLUDING REMARKS

WSNs data analysis can benefit greatly from the development of FDA. Functional data with spatial and temporal dependence is a new topic which offers the possibility of combining knowledge from spatial and spatio-temporal statistics with functional data analysis techniques. In this paper we used a flexible modeling approach for non-separable spatio-temporal data using tensor product smooths in Lee and Durbán (2011), which provides a flexible tool to predict and forecast spatio-temporal functional data. The use of the smooth modulation model with penalized splines (Eilers et al., 2008) allows us to decompose the functional data in terms of time trend and daily cyclical modulation components, with the advantage of considering a smaller number of basis functions, and hence gaining computational efficiency. Given the potentially huge amount of data collected by a WSN, dimensionality reduction techniques such as FPCA were used to find the main sources of variability and functional tests were applied to check the evidence of soil temperature differences between land types.

A possible extension of the proposed methodology is to allow for the principal components to vary according to the spatial longitude and latitude. That may help to improve the prediction in those locations close to the boundaries of each land type where residuals were found to be larger that in the central areas of each land types. Finally, the implementation of the proposed methods can be done in standard `R` packages for functional data analysis and smoothing making the approach very easy to apply and use.

Additional information and supplementary material for this article is available online at the journal's website.

ACKNOWLEDGEMENTS

Dae-Jin Lee and Peter Toscas were supported by an NIH grant for the Superfund Metal Mixtures, Biomarkers and Neurodevelopment project 1PA2ES016454-01A2. Most of this work was done during Dae-Jin Lee's Postdoctoral Fellowship at CSIRO. Dae-Jin was also supported by the Spanish Ministry of Economy and Competitiveness grant MTM2011-28285-C02-02 and also by the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013- 0323. The authors would also like to thank Dr Adrien Ickowicz for his helpful comments, Darren Moore and Brett Wood for their work in setting up and maintaining the sensor network, and the Sensors and Sequences Project in CSIRO for providing the data used in the paper.

REFERENCES

Benson, B., Bond, B., Hamilton, M., Monson, R., and Han, R. (2010). Perspectives on next-generation technology for environmental sensor networks. *Frontiers in Ecology and the Environment*, 8:183–200.

Buratti, C., Conti, A., Dardari, D., and Verdone, R. (2009). An overview on wireless sensor networks technology and evolution. *Sensors*, 9:6869–6897.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of cross-validation. *Numer. Math.*, 31:377–403.

Cressie, N. (1993). *Statistics for Spatial Data (Revised Edition).* John Wiley and Sons, Inc., New York.

Cressie, N. and Huang, H. C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(448):1330–1340.

de Boor, C. (1978). *A practical Guide to Splines.* Springer, Berlin.

Delicado, P., Giraldo, R., Comas, C., and Mateu, J. (2010). Statistics for spatial functional data: some recent contributions. *Environmetrics*, 21:224–239.

Eilers, P. H. C. and Marx, B. D. (2010). Splines, knots and penalties. *WIREs computational statistics*, 2(6):637–653.

Eilers, P. H. C., Currie, I. D., and Durbán, M. (2006). Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis*, 50(1):61–76.

Eilers, P. H. C., Gampe, J., Marx, B. D., and Rau, R. (2008). Modulation models for seasonal time series and incidence tables. *Statistics in Medicine*, 27:3430–3441.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Stat. Sci.*, 11:89–121.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional data analysis.* Springer, New York.

Fuentes, M., Chen, L., and Davis, J. M. (2007). A class of nonseparable and nonstationary spatial temporal covariance functions. *Environmetrics*, 19(5):487-507.

Giraldo, R., Delicado, P., and Mateu, J. (2011). Ordinary kriging for functional-valued spatial data. *Environmental and Ecological Statistics*, 18(3):411–426.

González-Manteiga, W. and Vieu, P. (2007). Editorial-introduction to the special issue on statistics for functional data. *Computational Statistics and Data Analysis*, 51:4788–4792.

Gu, C. (2002). *Smoothing Spline ANOVA Models.* Springer Series in Statistics. Springer-Verlag New York.

Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models.* Monographs on Statistics and Applied Probability. Chapman and Hall, London.

Hurvich, C. and Simonoff, J. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *J. R. Statist. Soc. B*, 60:271–293.

Hyndman, R. J. and Shang, H. L. (2013). *ftsa: Functional time series analysis*. R package version 3.9.

Hyndman, R. J. and Ullah, M. S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis*, 51(10):4942–4956.

Lee, D.-J. and Durbán, M. (2011). P-spline ANOVA-type interaction models for spatio-temporal smoothing. *Statistical modelling*, 11(1):49–69.

Lee, D.-J., Durbán, M., and Eilers, P.H.C. (2013). Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases *Computational Statistics and Data Analysis*, 61:22–37.

Nerini, D., Monestiez, P., and C., M. (2010). Cokriging for functional data. *Journal of Multivariate Analysis*, 101(2):409–418.

Pedesma, E. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30:683–691.

Ramsay, J. O. and Dalzell, C. J. (1991). Some tools for functional data analysis. *J. R. Statist. Soc. B*, 53:539–572.

Ramsay, J. O. and Silverman, B. (2005). *Functional data analysis*. Springer, New York, 2^{nd} ed edition.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, UK. ISBN: 0521785162.

Wahba, G. (1985). A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. *Annals of Statistics*, 13(4):1378–1402.

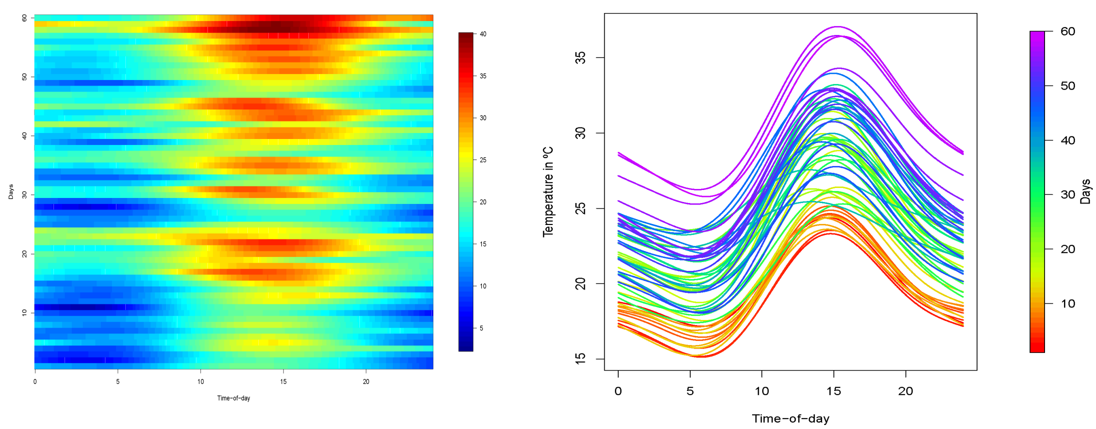
Wood, S. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Statist. Soc. B*, 73(1):3–36.

Wood, S. N. (2006). *Generalized Additive Models - An introduction with R*. Texts in Statistical Science. Chapman & Hall.

Yick, J., Biswanath, M., and Ghosal, D. (2008). Wireless sensor network survey. *Computer networks*, 52:2292–2330.

FIGURES

Figure 1. Wireless sensor nodes locations on the field



(a) Smoothed image plot of Sensor ID 1.

(b) Smoothed data of Sensor ID 1 by time-of-day.

Figure 2. Functional Data representation soil temperature data from Sensor ID1 using the smooth modulation model.

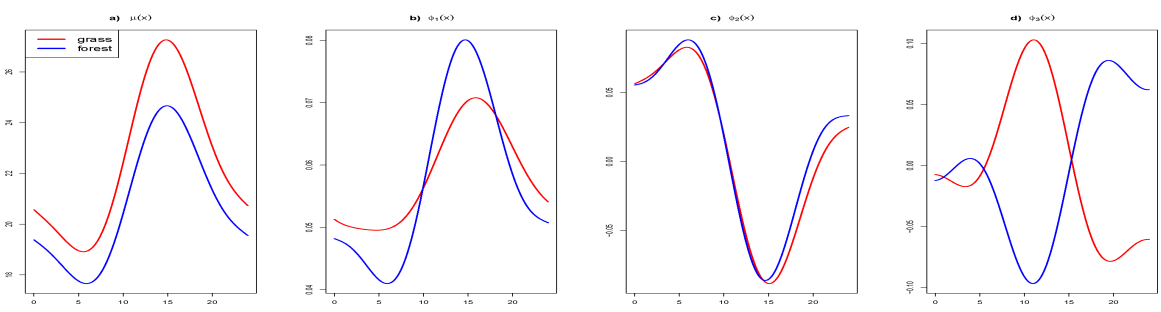
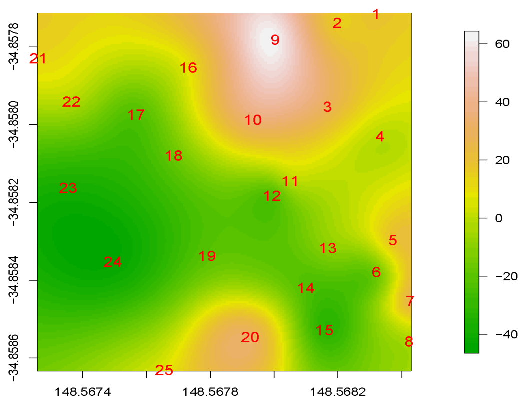
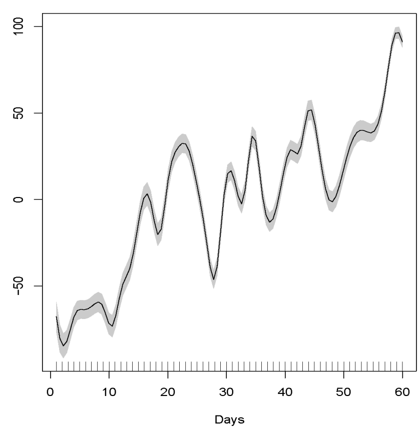


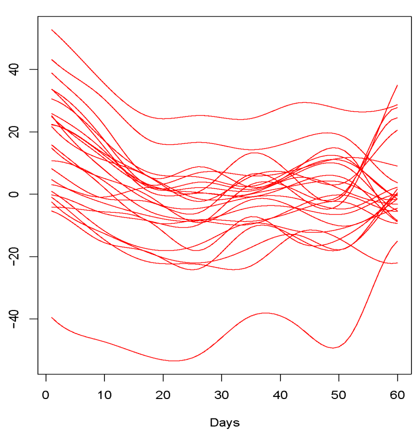
Figure 3. Functional mean, $\mu_1(x)$ and first three functional principal components and their associated principal component scores for soil temperature a sensor node.



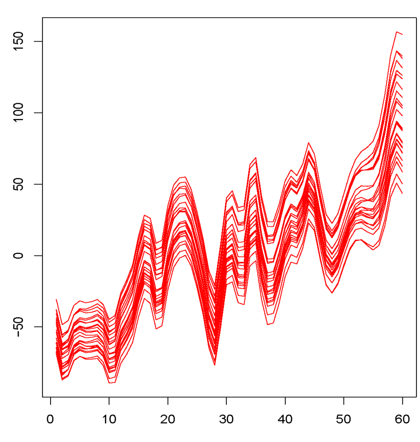
(a) Estimated spatial surface for grassland sensor nodes for the 1st PC score.



(b) Estimated time trend for grassland sensor nodes for the 1st PC score.

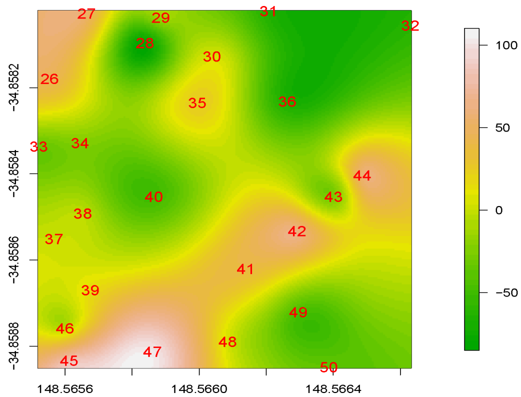


(c) Estimated space-time interaction for grassland sensor nodes for the 1st PC score.

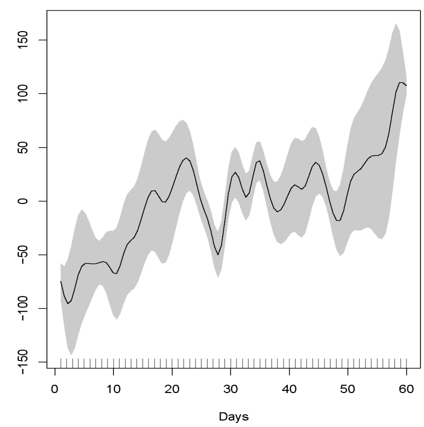


(d) Fitted spatio-temporal 1st PC score.

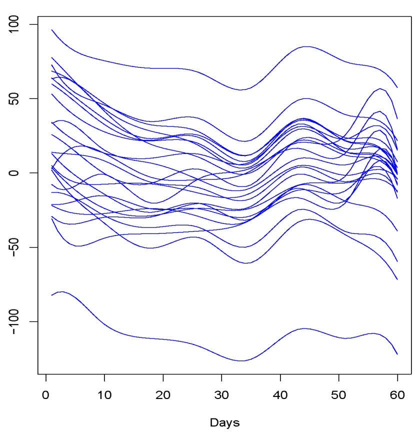
Figure 4. Smooth-ANOVA model decompostion of 1st PC scores of grassland locations.



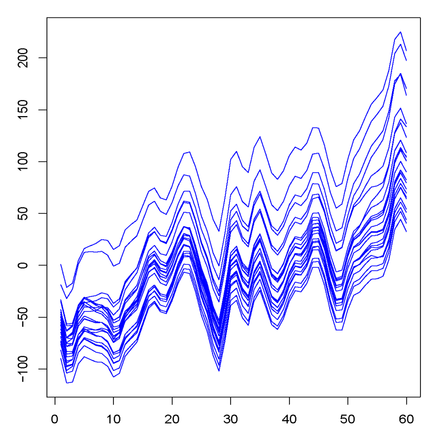
(a) Estimated spatial surface for forest sensor nodes for the 1st PC score.



(b) Estimated time trend for forest sensor nodes for the 1st PC score.

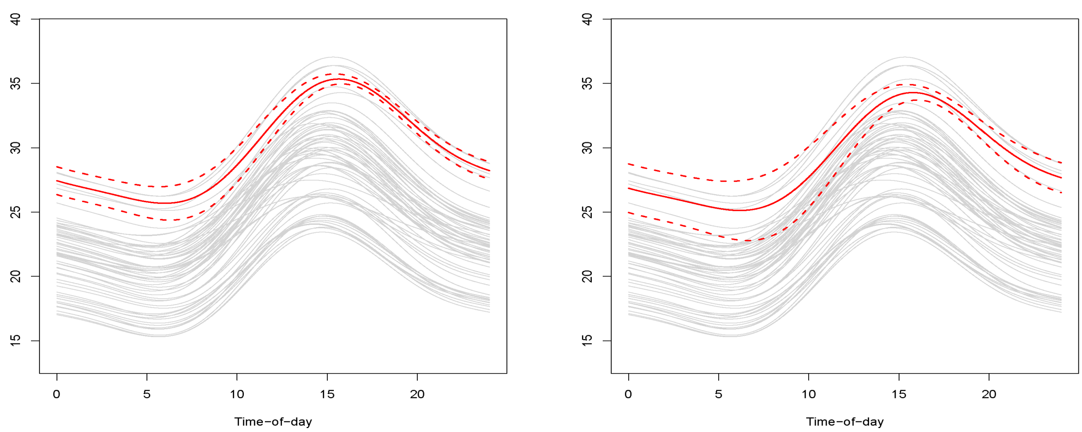


(c) Estimated space-time interaction for forest sensor nodes for the 1st PC score.



(d) Fitted spatio-temporal 1st PC score.

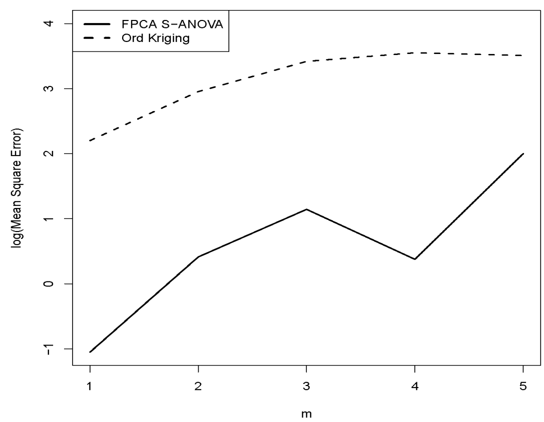
Figure 5. Smooth-ANOVA model decomposition of 1st PC scores of forest locations.



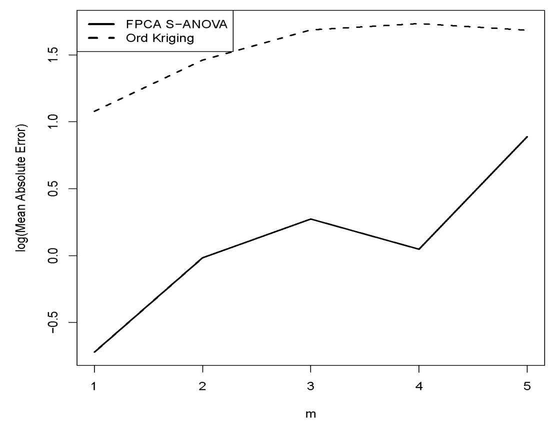
(a) $h = 1$

(b) $h = 2$

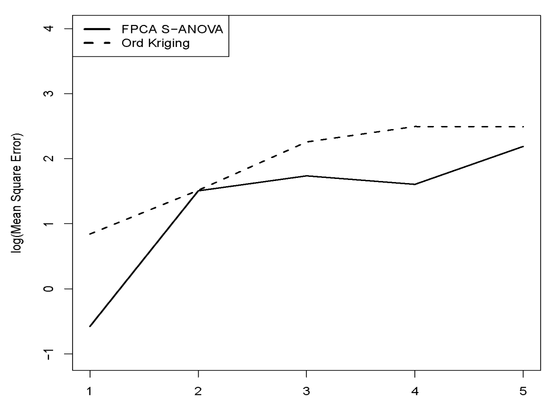
Figure 6. Soil temperature functional data forecast for the next two days $h = 1, 2$ for sensor node ID 1, with 95% prediction intervals.



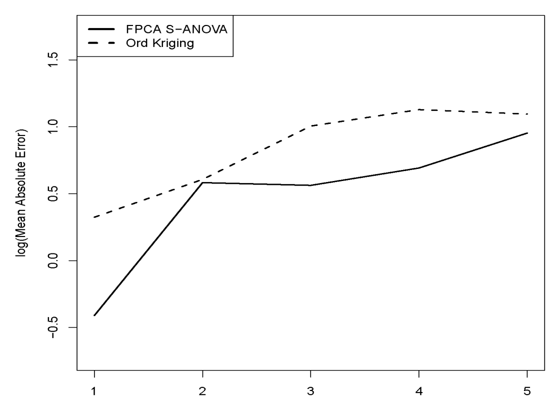
(a) out-of-sample MSE for sensor node ID 1 forecast.



(b) out-of-sample MAE for sensor node ID 1 forecast.



(c) out-of-sample MSE for sensor node ID 28 forecast.



(d) out-of-sample MAE for sensor node ID 28 forecast

Figure 7. Mean square and absolute errors of sensor nodes ID 1 amd ID 28. Fitting period is $60 - m$ where m=1,...,5.

TABLES

Table 1. Percentage of variability explained by functional principal components for sensor nodes in grassland and forest locations.

	1^{st} PC	2^{nd} PC	3^{rd} PC	Total
grassland	91.25%	6.44%	1.90%	99.59%
Forest	90.69%	7.16%	1.59%	99.44%